APPENDIX B

DID RANDOMIZATION PRODUCE EQUIVALENT GROUPS? A COMPARISON OF PROGRAM AND CONTROL GROUP CHARACTERISTICS

THIS PAGE INTENTIONALLY BLANK

A. INTRODUCTION

In theory, randomized experimental designs ensure that observed differences in outcomes between program and control groups can be attributed to the intervention under investigation, up to a known degree of statistical sampling error. This rigor is possible, however, only if the random assignment process generates program and control groups with similar characteristics at the time of random assignment. Thus, the benefits of a randomized design can be realized only if the random assignment process is implemented correctly.

We believe that the process used in the National Job Corps Study to randomly assign youths in the sample universe to the program or control groups was implemented correctly. MPR staff controlled the random assignment process, and random numbers generated from a computer were used to assign the youths.

In this appendix, we compare the characteristics of program and control group members to check that the random assignment process was implemented properly. Ideally, we would like to compare both observable and unobservable characteristics of sample members at random assignment. However, it is clearly not possible to compare unobservable characteristics. Thus, we will use data on a set of the observable measures and assume that if program and control group members are similar along observable dimensions, then they are also similar along unobservable dimensions.

Next, we discuss the data sources and methods used for the analysis. Finally, we discuss analysis results.

1. Data Sources and Methods

We use two data sources for the analysis. First, we use baseline interview data, which contain a rich set of variables for analysis. The disadvantage of using the baseline data, however, is that

interview responses to certain questions could have differed for program research and control group members, because the baseline interview was sometimes conducted *after* OA staff contacted youths about their research status.[1] Consequently, we also use data from the ETA-652 and ETA-652 Supplement forms. These data were collected *prior* to random assignment, so neither the quality of the data nor item responses should differ by research status if random assignment was implemented correctly.

We use standard statistical tests to assess the similarity of program research and control group members and examine the magnitude and patterns of any differences that exist. We use univariate t-tests to compare variable means for binary and continuous variables and chi-squared tests to compare distributions of categorical variables. All figures are calculated using sample weights, and the test statistics incorporate design effects due to unequal weighting of the data.[2]

In addition, we conduct a more formal multivariate analysis to test the hypothesis that key variable means and distributions are *jointly* similar. For this analysis, we estimate logit regression models where the probability an individual is a program research group member is regressed on a set of individual characteristics, and we use chi-squared tests to assess whether the coefficients on these explanatory variables are jointly significant.

This joint analysis is a more rigorous procedure than the univariate analysis for two main reasons. First, the univariate analysis is expected to produce significant test statistics for some of the large number of hypotheses by chance, even when the program research and control groups are identical. For example, if the hypothesis tests are conducted at the 10 percent level of significance,

---

[1]As discussed in Appendix C, it is likely that about one-quarter of respondents completed the baseline interview after they knew their research status.

[2]The test statistics using baseline interview data also incorporate design effects due to clustering caused by the selection of areas slated for in-person interviewing.

then we would expect that 10 percent of independent tests would be falsely rejected. The multivariate analysis avoids this multiple comparisons problem. Second, the joint test accounts for correlations across measures, whereas the univariate tests assume that the measures are independent.

It is common to specify a 5 percent significance level (Type I error) when conducting a statistical test for the hypothesis that a mean characteristic is the same for two independent samples. This implies that there is only a 5 percent chance that the null hypothesis will be rejected erroneously (that is, that the test will find a statistical difference when in fact there is none). This standard implies, however, that the researcher should assume that no differences between the two groups exist, unless there is strong evidence to the contrary. Consequently, this framework assumes that rejecting the hypothesis when it is true (the Type I error) is more serious than accepting the hypothesis when it is false (Type II error).

While this framework is appropriate when estimating program impacts using follow-up data, it is less appropriate when assessing the success of random assignment using baseline data. We believe that in our context, it is more appropriate to assume that differences across research groups do exist, unless there is strong evidence to the contrary (that is, when in doubt, we should assume random assignment was *not* properly implemented). Hence, in our case the Type II error is more serious than the Type I error.

Unfortunately, it is not possible to construct a formal test for the null hypothesis that a measure *differs* across the two research groups. Hence, our approach is to perform standard hypothesis tests, but to *increase* the Type I error, which thereby reduces the Type II error. Consequently, we use a 15 percent significance level to identify variables that differ by research status. Using this standard, if the true population proportion is 50 percent, we will report a significant difference at the 15 percent level if the difference between the sample proportions for program and control group

members exceeds 1.2 percentage points.[3] For a 10 percent proportion, the figure is .7 percentage points.

2. Analysis Results

Tables B.1 to B.12 display analysis results. The tables display variable distributions and means for control and program research group members, as well as p-values for testing differences across the two groups.

The program research and control groups have similar characteristics using statistics based on either program intake or baseline interview data. Only a small number of univariate tests are rejected at the 15 percent level of significance (that is, whose p-values are below .15), and there are only small differences across the two groups in those few variables for which significant differences exist. In addition, no patterns across the variables appear to differ. The multivariate regression analysis yields similar results (not shown). Finally, the joint tests from the regression models yield p-values of more than .70, using either baseline interview or ETA-652 data.

It is particularly important to note that the crime and drug use measures are similar by research status. As discussed in Appendix C, we were concerned in the design phase of the evaluation that the quality of these data items might differ for program and control group members. For example, we feared that program group members may have been more reluctant than control group members to report their criminal activities or drug use (which they did not report to OA counselors) for fear that this information would threaten their Job Corps eligibility. For this reason, the Supplemental ETA-652 forms included several questions on criminal involvement. However, we find few differences in the distribution of the measures by research status and no pattern in the reporting differences between the two groups. Thus, our analysis indicates that comparable baseline measures on crime and drug use can be obtained for both research groups through baseline interview data.

[3] The cutoff level would be 1.7 percentage points at the 5 percent level of significance.

TABLE B.1

COMPARISON OF THE DEMOGRAPHIC CHARACTERISTICS OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

Variable	Control Group	Program Group	P-Value for Testing Differences
Male	59.7	59.3	.56
Age at Application			.80
16 to 17	39.9	39.5	
18 to 19	32.3	32.2	
20 to 21	16.9	16.9	
22 to 24	10.9	11.4	
(Average age)	18.9	19.0	.43
Race/Ethnicity			.62
White, non-Hispanic	29.1	29.4	
Black, non-Hispanic	50.4	50.5	
Hispanic	15.1	14.8	
American Indian or Alaskan Native	3.6	3.2	
Asian or Pacific Islander	1.8	2.1	
Job Corps Region of Residence			.58
1	4.5	4.4	
2	7.7	7.2	
3	13.1	13.0	
4	22.7	23.4	
5	10.5	10.4	
6	14.7	15.2	
7/8	12.0	12.7	
9	9.6	8.9	
10	5.1	4.8	
Size of City of Residence			.79
Less than 2,500	8.3	8.8	
2,500 to 10,000	11.5	11.2	
10,000 to 50,000	19.2	19.7	
50,000 to 250,000	17.7	17.4	
250,000 or more	43.3	42.9	
PMSA or MSA Residence Status			.46
In PMSA	32.7	31.7	
In MSA	45.1	45.8	
In neither	22.2	22.5	
Legal Resident	98.9	98.6	.21
U.S Citizen[a]	94.1	94.4	.44

TABLE B.1 *(continued)*

Variable	Control Group	Program Group	P-Value for Testing Differences
Native Language[a]			.63
English	85.7	85.9	
Spanish	9.3	8.9	
Other	5.0	5.2	
Job Corps Application Date			.92
11/94 to 2/95	22.2	22.6	
3/95 to 6/95	29.2	29.1	
7/95 to 9/95	28.1	27.7	
10/95 to 12/95	20.5	20.6	
Random Assignment Date			.94
11/94 to 2/95	16.6	17.0	
3/95 to 6/95	29.2	29.2	
7/95 to 10/95	35.9	35.6	
11/95 to 2/96	18.3	18.3	
ETA-652/Baseline Interview Sample Size	**5,977/5,514**	**9,409/8,813**	

SOURCE: Data from ETA-652 forms and baseline interviews.

NOTE: All figures are calculated using sample weights.

[a]Data item comes from the baseline interview.

*Significantly different from zero at the .15 level, two-tailed test.

TABLE B.2

COMPARISON OF THE CHILDHOOD EXPERIENCES AND BACKGROUNDS OF PARENTS OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

Variable	Control Group	Program Research Group	P-Value for Testing Differences
Head of Household[a]			.93
Father	33.4	33.6	
Stepfather	5.2	5.3	
Mother	49.0	48.3	
Grandparent, aunt, or uncle	8.3	8.6	
Other	4.1	4.3	
Family Was on Welfare When Youth Was Growing Up			.48
Never	45.9	47.0	
Occasionally	21.8	21.1	
Half the time	11.6	11.1	
Most or all of the time	20.7	20.7	
Mother Had a High School Diploma[a]	67.3	66.3	.29
Father Had a High School Diploma[a]	70.5	69.4	.24
Sample Size	**5,514**	**8,813**	

SOURCE: Baseline interview data.

NOTE: All figures are calculated using sample weights.

[a]Data pertain to when the sample member was 14 years old.

*Significantly different from zero at the .15 level, two-tailed test.

TABLE B.3

COMPARISON OF FERTILITY AND LIVING ARRANGEMENTS OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

Variable	ETA-652 Data			Baseline Interview Data		
	Control Group	Program Research Group	P-Value for Testing Differences	Control Group	Program Research Group	P-Value for Testing Differences
Has Dependents (from 652 Data)/ Natural Children (from Baseline Data)	15.5	14.9	.34	17.9	18.1	.71
Number of Dependents/Natural Children[a]			.71			.85
1	63.8	65.4		70.0	68.6	
2	23.8	22.4		22.2	22.4	
3 or more	12.4	12.2		7.9	9.0	
Needs Child Care Plan If Enrolls in Job Corps	12.5	12.5	.89			
Household Membership						.46
Living with both parents				17.5	17.2	
Living with mother only				42.2	41.5	
Living with father only				5.9	6.0	
Living with another adult relative				12.2	11.8	
Living with adult nonrelatives				4.9	4.8	
Living with no other adults				17.3	18.7	
Family Status			.45			
Family head	13.1	13.8				
Family member	61.3	60.5				
Unrelated individual	25.6	25.8				
Average Family Size	3.2	3.2	.58			
In Public or Rent-Subsidized Housing				19.8	20.4	.40
Sample Size	**5,977**	**9,409**		**5,514**	**8,813**	

SOURCE: Data from ETA-652 forms and baseline interviews.

NOTE: All figures are calculated using sample weights.

[a] Data pertain to those with dependents/natural children.

* Significantly different from zero at the .15 level, two-tailed test.

TABLE B.4

COMPARISON OF THE SCHOOLING AND TRAINING EXPERIENCES
OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

	ETA-652 Data			Baseline Interview Data		
Variable	Control Group	Program Research Group	P-Value for Testing Differences	Control Group	Program Research Group	P-Value for Testing Differences
Highest Grade Completed			.68			.87
Below 9	14.9	15.6		14.1	14.6	
9 to 11	63.2	62.9		64.9	64.4	
12	21.2	20.8		18.7	18.7	
Above 12	0.8	0.8		2.3	2.3	
(Average grade)	10.1	10.0	.24	10.1	10.1	.62
Degrees, Diplomas, and Certificates						
High school diploma				18.2	17.8	.54
GED certificate				5.5	4.7	.03*
Vocational, technical, or trade diploma				2.0	2.2	.38
Other				3.5	3.7	.52
In School or Training in the Month Prior to Application to Job Corps				25.6	25.1	.50
Attended Any Education Program in Past Year				66.4	65.6	.30
Average Number of Months Enrolled in Education Programs in Past Year[a]				6.87	6.85	.84
Type of Most Recent Education Program[a]						.40
Elementary or middle school				3.0	2.6	
High school				60.2	58.9	
ABE program				3.8	3.5	
GED program				10.9	11.5	
Vocational, technical, or trade school				8.7	8.9	
Other				13.4	14.5	
Served in Military	1.2	1.0	.40			
Sample Size	**5,977**	**9,409**		**5,514**	**8,813**	

SOURCE: Data from ETA-652 forms and baseline interviews.

NOTE: All figures are calculated using sample weights.

[a] Data pertain to those who attended an education program in the year prior to random assignment.

* Significantly different from zero at the .15 level, two-tailed test.

TABLE B.5

COMPARISON OF THE EMPLOYMENT AND EARNINGS OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

Variable	Control Group	Program Research Group	P-Value for Testing Differences
Ever Had a Full-Time or Part-Time Job	78.8	80.0	.09*
Had a Job in the Past Year	64.0	64.9	.26
Number of Full-Time or Part-Time Jobs in the Past Year[a]			.68
1	51.8	51.2	
2	29.5	29.5	
3 or more	18.6	19.4	
(Average number)	1.7	1.8	.38
Months Employed in the Past Year[a]			.63
Less than 1	11.0	10.3	
1 to 3	23.2	23.7	
3 to 6	24.8	25.6	
6 to 9	18.7	18.9	
9 to 11	13.8	12.7	
12	8.6	8.8	
(Average number)	5.6	5.5	.41
Had a Job at Random Assignment	20.7	21.4	.32
Usual Weekly Hours of Work on Most Recent Job[a]			.64
1 to 19	13.5	13.1	
20 to 29	20.3	19.7	
30 or more	66.2	67.2	
(Average hours)	35.3	35.6	.32
Hourly Wage on Most Recent Job[a]			.75
Less than $4.25	9.1	9.5	
$4.25	19.9	20.3	
$4.25 to $5.00	21.5	20.8	
$5.00 to $6.50	37.3	36.6	
$6.50 or more	12.2	12.9	
(Average hourly wage in dollars)	5.1	5.1	.75
Earnings in the Past Year[a]			.59
Less than $1,000	19.6	18.5	
$1,000 to $2,500	22.5	23.6	
$2,500 to $5,000	22.7	23.3	
$5,000 to $10,000	23.9	23.6	
$10,000 or more	11.4	11.0	
(Average earnings in dollars)	4,626.2	4,584.6	.58
Sample Size	**5,514**	**8,813**	

TABLE B.5 *(continued)*

SOURCE: Baseline interview data.

NOTE: All figures are calculated using sample weights.

[a]Data pertain to those who had a job lasting more than two weeks during the year prior to random assignment.

*Significantly different from zero at the .15 level, two-tailed [or one-tailed] test.

TABLE B.6

COMPARISON OF THE WELFARE DEPENDENCE OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

	ETA-652 Data			Baseline Interview Data		
Variable	Control Group	Program Research Group	P-Value for Testing Differences	Control Group	Program Research Group	P-Value for Testing Differences
Received AFDC in the Past Year				31.6	31.5	.92
Received Food Stamps in the Past Year				44.6	43.7	.30
Received Other Public Assistance in the Past Year[a]				26.7	26.8	.90
Received Any Public Assistance in the Past Year				58.5	57.8	.45
Type of Welfare Received			.90			
AFDC	26.4	26.7				
Other types	16.4	16.6				
None	57.2	56.8				
Sample Size	**5,977**	**9,409**		**5,514**	**8,813**	

SOURCE: Data from ETA-652 forms and baseline interviews.

NOTE: All figures are calculated using sample weights. The welfare recipiency items on the baseline interview refer to income received either by the sample member or by the sample member's family in the year prior to random assignment.

[a]This assistance includes General Assistance, Supplementary Security Income (SSI), and Social Security Retirement, Disability, and Survivors Benefits (SSA).

*Significantly different from zero at the .15 level, two-tailed test.

TABLE B.7

COMPARISON OF TOTAL HOUSEHOLD AND PERSONAL INCOME OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS IN THE LAST CALENDAR YEAR
(Percentages)

Variable	Control Group	Program Research Group	P-Value for Testing Differences
Total Household Income			.66
Less than $3,000	25.4	25.9	
$3,000 to $6,000	20.8	19.5	
$6,000 to $9,000	10.8	11.3	
$9,000 to $18,000	24.4	24.6	
$18,000 or more	18.6	18.6	
(Average income in dollars)	8,969.4	8,986.8	.75
Total Personal Income			.06*
Less than $3,000	79.0	78.6	
$3,000 to $6,000	12.9	12.6	
$6,000 to $9,000	4.4	5.4	
$9,000 or more	3.7	3.5	
(Average income in dollars)	2,479.7	2,512.3	.37
Sample Size	**5,514**	**8,813**	

SOURCE: Baseline interview data.

NOTE: All figures are calculated using sample weights. Total household income includes the total income of all members of the respondent's household before taxes and other deductions and includes all sources of income. Total personal income includes the total income of the respondent before taxes and other deductions.

* Significantly different from zero at the .15 level, two-tailed test.

TABLE B.8

COMPARISON OF THE HEALTH STATUS OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

Variable	ETA-652 Data			Baseline Interview Data		
	Control Group	Program Research Group	P-Value for Testing Differences	Control Group	Program Research Group	P-Value for Testing Differences
Ever Had Any Serious Illnesses or Injuries	2.2	2.6	.12			
Ever Been Under the Care of Any Physical or Mental Health Care Provider in the Past Year	3.9	4.3	.24			
Have Any Health Conditions That Are Being Treated	3.4	3.3	.77			
Health Status						.39
Excellent				46.5	46.8	
Good				40.2	40.7	
Fair or Poor				13.3	12.5	
Has Physical or Emotional Problems That Limited the Amount of Work That Could Be Done				5.4	4.7	.04**
Covered by Health Insurance or Eligible for Medicaid	37.4	37.2	.89			
Sample Size	**5,977**	**9,409**		**5,514**	**8,813**	

SOURCE: Data from ETA-652 forms and baseline interviews.

NOTE: All figures are calculated using sample weights.

* Significantly different from zero at the .15 level, two-tailed test.

TABLE B.9

COMPARISON OF THE TOBACCO, ALCOHOL, AND ILLEGAL DRUG USE, AND DRUG TREATMENT OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

Variable	Control Group	Program Research Group	P-Value for Testing Differences
Smoked Cigarettes			
Ever	54.4	54.4	.97
In the past year	52.6	52.6	.93
Consumed Alcoholic Beverages			
Ever	57.7	59.7	.01*
In the past year	52.8	54.4	.05*
Smoked Marijuana or Hashish			
Ever	36.6	37.7	.18
In the past year	29.7	31.0	.07*
Summary of Drugs Ever Used			.31
Did not use drugs	62.4	61.4	
Used marijuana but not other drugs	28.7	29.0	
Used other drugs but not marijuana	0.9	0.9	
Used marijuana and other drugs	7.9	8.7	
Ever in a Drug or Alcohol Treatment Program	5.5	4.7	.10*
Sample Size	**5,514**	**8,813**	

SOURCE: Baseline interview data.

NOTE: All figures are calculated using sample weights.

* Significantly different from zero at the .15 level, two-tailed test.

TABLE B.10

COMPARISON OF THE ARREST EXPERIENCES OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

Variable	ETA-652 Supplemental Data			Baseline Interview Data		
	Control Group	Program Research Group	P-Value for Testing Differences	Control Group	Program Research Group	P-Value for Testing Differences
Arrested in Past Three Years, Other than for Minor Traffic Violations	12.1	11.9	.80	22.9	22.7	.71
Ever Arrested or Charged with a Delinquency or Criminal Complaint				27.0	25.9	.79
Number of Times Ever Arrested[a]						.69
1				61.5	60.5	
2				20.8	22.2	
3				8.6	8.9	
4 or more				9.2	8.4	
Number of Months Since Most Recently Arrested[a]						.95
Less than 12				48.9	48.4	
12 to 24				24.0	24.5	
24 or more				27.1	27.1	
Most Serious Charge for Which Arrested[a]						.01*
Murder or assault				8.3	10.0	
Robbery				3.3	2.7	
Burglary				8.2	8.4	
Larceny, vehicle theft, or other property crimes				29.7	33.6	
Drug law violations				8.0	7.3	
Other personal crimes[b]				14.4	2.9	
Other miscellaneous crimes[c]				28.0	27.4	
Sample Size	**5,977**	**9,409**		**5,514**	**8,813**	

SOURCE: Data from ETA-652 Supplemental forms and baseline interviews.

NOTE: All figures are calculated using sample weights.

TABLE B.10 *(continued)*

---

[a] Data pertain to those who were ever arrested.

[b] "Other personal crimes" include simple assault, family offenses, sex offenses other than rape, and fighting.

[c] "Other miscellaneous crimes" include disorderly conduct, liquor law violations, gambling, loitering, being a Peeping Tom, trespassing, having an outstanding warrant, pornography-related offenses, obstruction of justice, truancy, and motor vehicle violations.

* Significantly different from zero at the .15 level, two-tailed test.

TABLE B.11

COMPARISON OF THE CONVICTION EXPERIENCES OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

	ETA-652 Data			Baseline Interview Data		
Variable	Control Group	Program Research Group	P-Value for Testing Differences	Control Group	Program Research Group	P-Value for Testing Differences
Ever Convicted or Adjudged Delinquent[a]	6.1	5.9	.69	17.0	16.3	.84
Number of Times Convicted[b]						.33
1				54.5	57.5	
2				29.5	26.6	
3 or more				15.9	15.8	
Ever Made a Deal or Copped a Plea				6.3	5.3	.02**
Ever Served Time in Jail				8.3	7.7	.39
Ever Put on Probation or Parole				12.0	11.5	.99
Currently on Probation or Parole				3.9	4.0	.46
Sample Size	**5,977**	**9,409**		**5,514**	**8,813**	

SOURCE: Data from ETA-652 forms and the baseline interview.

NOTE: All figures are calculated using sample weights.

[a]The data item for the baseline interview also includes those who ever pled guilty.

[b]Data pertain to those who were ever convicted, pled guilty, or adjudged delinquent.

*Significantly different from zero at the .15 level, two-tailed test.

TABLE B.12

COMPARISON OF THE ANTICIPATED PROGRAM ENROLLMENT INFORMATION OF CONTROL AND PROGRAM RESEARCH GROUP MEMBERS
(Percentages)

Variable	Control Group	Program Research Group	P-Value for Testing Differences
Designated for a Nonresidential Slot	13 9	13 7	0 78
Designated for a CCC Center	14 7	15 1	0 53
1995 Performance Ranking of Designated Center (Quartlies)[a]			0 85
Lowest	27 2	27 4	
Second-lowest	28 2	28 7	
Second-highest	24 9	24 4	
Highest	19 7	19 5	
Size of Designated Center in 1995 (Slots)[a]			0 62
Small (225 slots or less)	19 8	20	
Medium small (225 to 495 slots)	45 4	45 3	
Medium large (496 to 735 slots)	20	19 3	
Large (more than 735 slots)	14 8	15 4	
Estimated Number of Weeks from Application Interview Until Arrival at Center			0 23
Less than 2	11 4	11 9	
2 to 3	11 5	11 8	
3 to 4	37 2	36 5	
4 to 8	30 4	29 4	
8 or more	9 4	10 4	
(Average weeks)	5 8	5 8	0 5
Likelihood of Enrolling in a Center			0 46
Very likely	83 3	83	
Somewhat likely	15 5	15 6	
Somewhat unlikely	1	1 1	
Very unlikely	0 2	0 4	
Sample Size	**5977**	**9409**	

SOURCE: Data from ETA-652 Supplemental forms.

NOTE: All figures are calculated using sample weights.

[a] Figures are obtained using data on OA counselor projections about the centers that youths were likely to attend.

* Significantly different from zero at the .15 level, two-tailed test.

THIS PAGE INTENTIONALLY BLANK